



## **Supplemental User Note 2: Record Lengths**

Numerical Identification Files (NUMIDENT)  
TR-0047-2013-0013  
Records of the Social Security Administration (Record Group 47)

In reviewing the layouts and data files, staff noted that the record length in the data files does not always match what is in the layout.

In particular, the claim records contain 498 characters with the end of record marker (carriage return and line feed) in columns 499 and 500. Thus the records do not include the field NUMI-FLAG-DELETE-IND (column 500). The agency did not indicate why they did not include that field in the transfer.

The death records contain 487 characters with the carriage return and line feed in columns 488 and 489. So instead of filler in columns 488 to 500, it appears the agency put the end of record markers after the last column with valid data.

The SS-5 records contain fixed-length fields, but the record length varies depending on the number of the fields in the record and may be less than 500 characters. Many of the SS-5 records contain 498 characters with the carriage return and line feed in columns 499 and 500. Thus these records do not include the fields NUMI-ENTRY-SUBCODE (column 499) and NUMI-FLAG-DELETE-IND (column 500).

However, some records do have data in the field NUMI-ENTRY-SUBCODE (column 499) and thus have a record length of 499 characters with the carriage return and line feed in columns 500 and 501. Below is a list of the number of records per file with valid codes in the NUMI-ENTRY-SUBCODE.

File	Number of records with NUMI-ENTRY- SUBCODE field
NARA.S01.R130915.SS5	2562
NARA.S02.R130915.SS5	2472
NARA.S03.R130915.SS5	2461

NATIONAL ARCHIVES *and*  
RECORDS ADMINISTRATION

8601 ADELPHI ROAD  
COLLEGE PARK, MD 20740-6001  
[www.archives.gov](http://www.archives.gov)

NARA.S04.R130915.SS5	2447
NARA.S05.R130915.SS5	2559
NARA.S06.R130915.SS5	2498
NARA.S07.R130915.SS5	2463
NARA.S08.R130915.SS5	2520
NARA.S09.R130915.SS5	2521
NARA.S10.R130915.SS5	2082
NARA.S11.R130915.SS5	2488
NARA.S12.R130915.SS5	2542
NARA.S13.R130915.SS5	2359
NARA.S14.R130915.SS5	2601
NARA.S15.R130915.SS5	2579
NARA.S16.R130915.SS5	2500
NARA.S17.R130915.SS5	2453
NARA.S18.R130915.SS5	2399
NARA.S19.R130915.SS5	2540
NARA.S20.R130915.SS5	2526

It is possible some records include data in both the fields NUMI-ENTRY-SUBCODE (column 499) and NUMI-FLAG-DELETE-IND (column 500) so match up with the record length in the layout.

Staff also noticed at least one SS-5 record with the end of record marker in columns 337 and 338. It is possible there are other records with fewer characters than indicated in the layout.

Lynn Goodsell, Reference Branch Chief  
Electronic Records Division (RDE)  
March 3, 2015



**Supplemental User Note 3:  
Death Records Public Use Versions**

Numerical Identification Files (NUMIDENT)  
TR-0047-2013-0013  
Records of the Social Security Administration (Record Group 47)

In reviewing the Death Records Files, staff discovered that the agency occasionally captured information other than death certificates in the Death Certificate Number field (NUMI\_DTH\_CERTIFICATE\_NUM) and that the Other Number field (NUMI-INDIVIDUALS-OTHER-NUM) may potentially contain social security numbers.

The Death Certificate Number field is only populated in about 7% of the records. While some of the records appear to contain death certificate numbers in this field, other records have data entry notes, funeral home names, and phone numbers. Since some of the records were created in the past 20 years and the phone numbers in the Death Certificate Number field may be for private individuals, the National Archives determined this information is considered an invasion of personal privacy and exempt from disclosure under the Freedom of Information Act (FOIA)(b)(6).

Even though not all the records had phone numbers in the Death Certificate Number field, staff had no automated means to withhold just the phone numbers. Instead, staff created public use versions that masked the values in the Death Certificate Number field in the 7% of the records that contained information in that field. The information in the Death Certificate Number field was masked with an "X."

The Other Number field contains the social security number of the decedent in 71% of the records and is blank in 13% of the records. For the remaining records, it is possible the Other Number may contain a social security number of a living individual, which is exempt from disclosure under FOIA(b)(6).

In the public use version, staff masked the values in the Other Number field that did not match a social security number of a deceased individual listed in the NUMIDENT Death Files. NARA masked the values in this field with "XXXXXXXXXX" in close to 9% of the records.

Lynn Goodsell, Reference Branch Chief  
Electronic Records Division (RDE)  
April 12, 2017

NATIONAL ARCHIVES *and*  
RECORDS ADMINISTRATION

8601 ADELPHI ROAD  
COLLEGE PARK, MD 20740-6001  
[www.archives.gov](http://www.archives.gov)



**Supplemental User Note 4:  
Application (SS-5) Records Public Use Versions**

Numerical Identification Files (NUMIDENT)  
TR-0047-2013-0013  
Records of the Social Security Administration (Record Group 47)

**Public Use Versions**

In reviewing the Application (SS-5) records files, staff discovered that a relatively small number of records were of potentially living individuals. Since the files were to only contain application records of deceased individuals, the records of potentially living individuals are exempt from disclosure under the Freedom of Information Act (FOIA)(b)(6) because of the personal information in the records. NARA created public use versions of the files.

Staff identified records of potentially living individuals by comparing the SS-5 records to the NUMIDENT Death records received in the same transfer. Staff masked those SS-5 records that did not have a corresponding Death record for that individual and the individual was born in or after 1908. NARA did not mask records of individuals born over 110 years ago (born prior to 1908), regardless if there was a corresponding Death record or not. NARA masked the records of potentially living individuals by replacing the contents of the records with "Z"s. In the masked records, all fields are populated with Z, except for the two field. In the NUMI-NOTIFY field (positions 358-359), there is only a single Z so to differentiate it from the coded value ZZ. In the filler field at positions 422-496, there is only a single Z.

In the public use versions of the files, a total of 60,213 records are masked.

Number of records masked per file:

File 1	2,990
File 2	2,547
File 3	3,217
File 4	2,967
File 5	2,976
File 6	3,123
File 7	3,177
File 8	2,915
File 9	3,108
File 10	2,951

File 11	2,944
File 12	3,097
File 13	2,958
File 14	2,969
File 15	2,951
File 16	3,225
File 17	2,985
File 18	3,101
File 19	3,058
File 20	2,954

NATIONAL ARCHIVES *and*  
RECORDS ADMINISTRATION

8601 ADELPHI ROAD  
COLLEGE PARK, MD 20740-6001

[www.archives.gov](http://www.archives.gov)

### Converting asterisk (\*) characters

For several fields, the agency used the asterisk (\*) character as a code. Since \* is used by many systems for wildcard searching, NARA staff converted these characters as part of creating the public use version of the SS-5 files.

In the fields that indicate a name or place of birth exceeds the field size (the field name has “OFLO”), NARA converted the \* to the capital letter O (for overflow).

In the NUMI-DOB-CHANGE-IND field, NARA converted the \* to the capital letter C (for corrected).

In the NUMI-POB-FOREIGN-IND field, NARA converted the \* to the lower-case letter f (for foreign). The lower-case f was used so to distinguish it from the code values used in the NUMI-POB-STATE-COUNTRY field.

As an update to the example in Supplemental User Note 1, if NUMI-POB-STATE-COUNTRY field contains CA and the NUMI-POB-FOREIGN-IND contains f, then the meaning of CA is Canada. If NUMI-POB-STATE-COUNTRY field contains CA and the NUMI-POB-FOREIGN-IND is blank, then the meaning of CA is California.

Stray \* characters (i.e. dirty data) in the NUMI-POB-STATE-COUNTRY field were also converted to lower-case letter f.

The number of \* converted per field:

Field	Number * converted	Converted to
NUMI-NH-NAME-FIRST-OFLO	458	O
NUMI-NH-NAME-MIDDLE-OFLO	31,074	O
NUMI-NH-NAME-LAST-OFLO	1,287	O
NUMI-MTH-NAME-FIRST-OFLO	3	O
NUMI-MTH-NAME-MIDDLE-OFLO	164	O
NUMI-MTH-NAME-LAST-OFLO	29	O
NUMI-FTH-NAME-FIRST-OFLO	9	O
NUMI-FTH-NAME-MIDDLE-OFLO	77	O
NUMI-FTH-NAME-LAST-OFLO	15	O
NUMI-POB-CITY-OFLO	24,595,578	O
NUMI-POB-STATE-COUNTRY	7,383	f
NUMI-POB-FOREIGN-IND	5,491,807	f
NUMI-OTH-NAME-FIRST-OFLO	4	O
NUMI-OTH-NAME-MIDDLE-OFLO	215	O
NUMI-OTH-NAME-LAST-OFLO	54	O
NUMI-SIG-NAME-FIRST-OFLO	2	O
NUMI-SIG-NAME-MIDDLE-OFLO	57	O
NUMI-SIG-NAME-LAST-OFLO	51	O
NUMI-DOB-CHANGE-IND	719,134	C

#### Other modifications

As part of creating the public use versions, staff also removed leading spaces from the name fields, such as NUMI-NH-NAME-FIRST, NUMI-NH-NAME-LAST, etc.

Staff found and removed stray binary characters in File 2, position 497 in approximately 7,050 records.

Staff replaced binary characters in these files with a space:

File 4, position 234 in line 2230589 (SSN 415276819)

File 19, position 235 in line 178479 (SSN 044301894)

File 19, position 274 in line 1714324 (SSN 326387490)

File 19, position 273 in line 2894825 (SSN 498091292)

#### Record length

In the public use versions, the record length is fixed at 500 characters. There is a line feed at the end of each record for a total length of 501 characters. The information about records length in User Note 2 is not applicable to the public use files

Lynn Goodsell, Reference Branch Chief  
Electronic Records Division (RDE)  
August 29, 2018



**Supplemental User Note 5:  
Number of Records and Unique SSNs**

Numerical Identification Files (NUMIDENT)  
TR-0047-2013-0013  
Records of the Social Security Administration (Record Group 47)

The number of records in each set of NUMIDENT records does not necessarily reflect the number of unique social security numbers (SSNs) reflected in the records. In addition, there may not be a record for a SSN in all three sets. For example, there may be a record for a SSN in the Application records, but not in the Death records.

Application Records

There are 72,180,729 Application (SS-5) records, but only 40,873,455 unique SSNs reflected in the records. In the public use versions of the Application records, there are 40,870,456 unique SSNs since in the masked records the SSN is the same (ZZZZZZZZZ).

For the Application (SS-5) records, there is a record each time an individual submitted a new application as a result of a name change, correction, or other reason. It is also possible that the Social Security Administration created additional records as part of updating information. As a result, there may be multiple records for a single SSN. While in most cases there are only two to three records for a single SSN, staff found instances of up to 30 or more records for a single SSN. The NUMI-ENTRY-CODE and NUMI-CYCLE-DATE may help in distinguishing multiple records for a single SSN.

There are 5,957,896 SSNs in the Application records that do not appear in the Death records.

Death Records

For the Death records, there are 49,459,293 records and each record is for a unique SSN.

Claim Records

For the Claim records, there are 25,230,486 records and 25,140,847 unique SSNs reflected in the records.

There are 22,906,070 SSNs in the Claims records that do not appear in the Application records.  
There are 8,656,764 SSNs in the Claims records that do not appear in the Death records.



## **Supplemental User Note 6: Claim Records Public Use Versions**

Numerical Identification Files (NUMIDENT)  
TR-0047-2013-0013  
Records of the Social Security Administration (Record Group 47)

### Public Use Versions

To verify that the Claim records files only contained records of deceased individuals, staff compared the social security numbers (SSN) in the Claim records to those in the Death records and released Application records (see Supplemental User Notes 4). Staff found that 8,585,914 SSNs in the Claims records files are not in the Death and released Application records. Staff further filtered these records based on a date of birth in or after 1908 or no birth year. This process resulted in 2,229 records that did not have a corresponding SSN in the Death or released Application records and had a date of birth in or after 1908 or no birth year.

Since NARA does not have other means to systematically verify if the individuals in these 2,229 records are deceased, there is a possibility some of these records may be for potentially living individuals. The records of potentially living individuals are exempt from disclosure under the Freedom of Information Act (FOIA)(b)(6) because of the personal information in the records.

NARA created public use versions of the files that mask the 2,229 records that did not have a corresponding SSN in the Death or released Application records for that individual and had a date of birth in or after 1908 or no birth year. NARA masked the records by replacing the contents of the records with "Z"s. In the masked records, all fields are populated with Z, except for the filler field at positions 306-497 where there is only a single Z.

Number of records masked per file:

File 1	105
File 2	105
File 3	105
File 4	99
File 5	117
File 6	105
File 7	107

File 8	115
File 9	132
File 10	118
File 11	122
File 12	105
File 13	114
File 14	112

File 15	92
File 16	114
File 17	124
File 18	128
File 19	120
File 20	90

NATIONAL ARCHIVES *and*  
RECORDS ADMINISTRATION

8601 ADELPHI ROAD  
COLLEGE PARK, MD 20740-6001

[www.archives.gov](http://www.archives.gov)



### Converting asterisk (\*) characters

For several fields, the agency used the asterisk (\*) character as a code. Since \* is used by many systems for wildcard searching, NARA staff converted these characters as part of creating the public use version of the Claim files.

In the fields that indicate a name or place of birth exceeds the field size (the field name has “OFLO”), NARA converted the \* to the capital letter O (for overflow). However, many of the overflow fields are all blank or do not contain an \*. NARA only converted the \* in the NUMI-NH-NAME-FIRST-OFLO, NUMI-NH-NAME-LAST-OFLO, and NUMI-POB-CITY-OFLO fields.

In the NUMI-POB-FOREIGN-IND field, NARA converted the \* to the lower-case letter f (for foreign). The lower-case f was used so to distinguish it from the code values used in the NUMI-POB-STATE-COUNTRY field.

As an update to the example in Supplemental User Note 1, if NUMI-POB-STATE-COUNTRY field contains CA and the NUMI-POB-FOREIGN-IND contains f, then the meaning of CA is Canada. If NUMI-POB-STATE-COUNTRY field contains CA and the NUMI-POB-FOREIGN-IND is blank, then the meaning of CA is California.

The number of \* converted per field (including now masked records):

Field	Number * converted	Converted to
NUMI-NH-NAME-FIRST-OFLO	51	O
NUMI-NH-NAME-LAST-OFLO	204	O
NUMI-POB-CITY-OFLO	382,195	O
NUMI-POB-FOREIGN-IND	1,142,092	f

### Other modifications

As part of creating the public use versions, staff also removed leading spaces from the NUMI-NH-NAME-LAST field.

Staff replaced binary characters in these files with a space:

File 12, position 246 in line 1128035 (SSN 541160855)

File 16, position 237 in line 1095754 (SSN 527128676)

File 20, position 239 in line 902056 (SSN 450866595)

### Record length

In the public use versions, the record length is fixed at 499 characters. There is a line feed at the end of each record for a total length of 500 characters. The records do not include the field NUMI-FLAG-DELETE-IND (column 500). The information about records length in User Note 2 is not applicable to the public use files.

Lynn Goodsell, Reference Branch Chief  
Electronic Records Division (RDE)  
August 16, 2019